

Library Linked Data Incubator Group Review of the Final Report

F. Tim Knight
Associate Librarian
Osgoode Hall Law School Library
York University
December 15, 2011

These slides are a distillation of the Final Report issued by the W3C Library Linked Data Incubator Group.

<<http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/>>

This presentation was delivered on 15Dec2011 at a meeting of the Linked Data Interest Group at York University and intended to facilitate discussion.

FTK

Main Points of the Report

- Benefits of library linked data
- Current issues of traditional library data
- Library linked data initiatives
- Legal rights over library data
- Recommendations for next steps

Defining “Library linked data”

Terms Used in the Report

- Library
- Library Data
- Linked data
- Open Data
- Library linked data

Defining “Library linked data”

Library

- Cultural heritage and memory institutions
(includes libraries, museums, archives)
- A collection, a place, an agent
(curator/administrator)

Defining “Library linked data”

Library Data

- Digital information produced/curated by libraries
- Describes resources/aids their discovery
- Typically three types:
 - **datasets** (e.g. British National Bibliography; Open Library)
 - **metadata element sets** (e.g. DCMI Metadata Elements; RDA elements; SKOS)
 - **value vocabularies** (e.g. LCSH; VIAF; DDC)

Defining “Library linked data”

linked data

- Data designed to facilitate links between:
 - datasets
 - metadata element sets
 - value vocabularies
- Expressed using standards (e.g. RDF)
- Relationships for navigation or integrating information from multiple sources

Defining “Library linked data”

Open Data

- “Legal interoperability”
- Freely usable, reusable and redistributable
- Requirements limited to “attribute and share alike”

Defining “Library linked data”

Library linked data

- Any type of Library Data expressed as linked data

Benefits of linked data

- Libraries can increase the value of their data beyond the individual sum of their sources
- Resources can be cited across a broader range of data sources
- Metadata descriptions become more accessible
- Capitalizes on the Library's ability to provide trusted metadata for resources of long-term cultural importance

Benefits of linked data

- Data distributed as statements rather than complete records
- “Graph-based ecosystem”
 - Individual statements about resources
 - Collected and aggregated into a “global graph”
- Important connections made from previously unknown sources
- Redundancy of metadata descriptions reduced

Benefits of linked data

- Linked data describes the meaning of data ("semantics") separately from specific data structures ("syntax" or "formats")
- Linked data retains its meaning across changes of format
- Linked data is more durable and robust than metadata formats that depend on a particular data structure

Benefits of linked data

Benefits to Researchers/Information Users

- Structured data will enhance the current Web of documents
- Improved capabilities for resource discovery and data use across library and non-library resources
- URI links provide a browsable global information graph - “toURIs”

Benefits of linked data

Benefits to Researchers/Information Users

- Improve library visibility through search engine optimization (SEO)
- Library data integrated into research documents and bibliographies
- Interdisciplinary research enhanced through links across multiple information domains
- Document/dataset links improves transparency of research and peer validation of results

Benefits of linked data

Benefits to Organizations

- Different data about the same resource can be produced by different actors and aggregated into a single graph
- Use of linked data technology and standards increases choice of vendors and allows interaction with a larger pool of developers
- "Cloud-based" approach to managing cultural information

Benefits of linked data

Benefits to Organizations

- Openness of data should be considered an opportunity rather than a threat
- Clarification of the licensing conditions of descriptive metadata facilitates reuse and improves institutional visibility
- “The coolest thing to do to your data will be thought of by someone else”

Benefits of linked data

Benefits to Librarians, Archivists and Curators

- Creation of an open, global pool of shared data
- Data can be used and re-used to describe resources
- Reduction of redundant effort compared with current cataloguing processes
- Cataloguers will concentrate their effort on their domain of local expertise

Benefits of linked data

Benefits to Librarians, Archivists and Curators

- Creation of an open, global pool of shared data
- Data can be used and re-used to describe resources
- Reduction of redundant effort compared with current cataloguing processes
- Cataloguers will concentrate their effort on their domain of local expertise

Benefits of linked data

Benefits to Developers and Vendors

- By leveraging RDF and HTTP development not tied to library centric formats and protocols, e.g. MARC, Z39.50
- Marketing of products outside of library world
- Larger developer community to support information technology in libraries

Current Issues of Traditional Library Data

- Library data in databases is not integrated with data sources on the Web
- Library standards designed for the library community
- Library data expressed in natural-language text
- Difference in terminology between library and linked data communities
- Changes in library technology depends on library system development by library vendors

Library Linked Data Initiatives

- Focus has been primarily on value vocabularies and metadata element sets rather than datasets
- Few bibliographic datasets
- Lacking metadata for journal articles, citations and circulation data
- Challenges: licensing; data modelling; legacy data; multiple user communities

Library Linked Data Initiatives

- Increased interest and activity but largely “prototypes” so long term stability is unsure
- Many national libraries have started work in this area which will contribute to long term stability
- Need to develop connections across datasets
- Need for long term support; better communication between developers; mature tools to produce semantic links

Legal Rights Over Library Data

- Complexity of rights ownership of library data
- Library data can have restricted usage based on local policies, contracts, and other conditions
- Hinders ability to provide open data
- Differences of rights between countries affects ability to collaborate internationally

Legal Rights Over Library Data

- High degree of data sharing over last 50 years
- Records, copied, modified then reshared in regional bibliographic utilities
- *“Assigning legally sound intellectual property rights between relevant agents and agencies is difficult, and the lack of certainty hinders data sharing in a community that is necessarily cautious on legal matters.”*

Legal Rights Over Library Data

- Libraries that have not participated in sharing may consider records as business assets
- Reluctant to publish as open linked data
- May be willing to provide brief or incomplete records which lowers effectiveness of the data

Recommendations

- Generally if libraries make their data available for use as Linked Data it will integrate library data into web of information, provide greater visibility and bring library services to information seekers
- Libraries can lead by:
 - Managing resources for current use and long term preservation
 - Describing resources based on standardized rules
 - Responding to the needs of information seekers

Recommendations

For Library Leadership

- Identify sets of data as possible candidates for early exposure as Linked Data
 - identify high-priority, low-effort Linked Data projects, e.g. authority files and controlled vocabulary lists
- Foster discussions about Open Data and rights
 - seek agreement with owners about rights and licensing at the level of library consortia or even on a national or international scale

Recommendations

For Standards Bodies and Participants

- Increase library participation in Semantic Web standardization
- Develop library data standards that are compatible with Linked Data
- Develop and disseminate best-practice design patterns tailored to library Linked Data

Recommendations

For Data and Systems Designers

- Design and test user services based on Linked Data capabilities
- Create URIs for the items in library datasets
- Develop policies for managing Linked Data vocabularies and their URIs
- Express library data by re-using or mapping to existing Linked Data vocabularies

Recommendations

For Librarians and Archivists

- Preserve Linked Data element sets and value vocabularies
- Apply library experience in curation and long-term preservation to Linked Data datasets